



PCT371

RAW SEQUENCE LISTING

DATE: 03/05/2004

PATENT APPLICATION: US/09/786,377A

TIME: 12:05:41

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03052004\I786377A.raw

3 <110> APPLICANT: Loma Linda University
 4 Szalay, Aladar A.
 5 Wang, Yubao
 6 Gefu, Wang-Pruski
 8 <120> TITLE OF INVENTION: Method for Studying Protein Interactions in Vivo
 10 <130> FILE REFERENCE: 11785-3
 12 <140> CURRENT APPLICATION NUMBER: US 09/786377A
 13 <141> CURRENT FILING DATE: 1999-09-02
 15 <150> PRIOR APPLICATION NUMBER: US 60/135,835
 16 <151> PRIOR FILING DATE: 1999-05-24
 18 <150> PRIOR APPLICATION NUMBER: US 60/099,068
 19 <151> PRIOR FILING DATE: 1998-09-03
 21 <150> PRIOR APPLICATION NUMBER: PCT/US99/20207
 22 <151> PRIOR FILING DATE: 1999-09-02
 24 <160> NUMBER OF SEQ ID NOS: 10
 26 <170> SOFTWARE: PatentIn version 3.1
 28 <210> SEQ ID NO: 1
 29 <211> LENGTH: 918
 30 <212> TYPE: DNA
 31 <213> ORGANISM: Homo sapiens
 33 <220> FEATURE:
 34 <221> NAME/KEY: CDS
 35 <222> LOCATION: (3)..(719)
 36 <223> OTHER INFORMATION:

ENTERED

W--> 39 <400> 1
 40 tg tgc ccc cac agg ctg ctg cca cgc ctg ctg ctg cta gct ctg 47
 41 Cys Pro His Arg Leu Leu Pro Pro Leu Leu Leu Leu Ala Leu
 42 1 5 10 15
 44 ctg ctc gct gcc agc cca gga ggc gcc ttg gcg cgg tgc cca ggc tgc 95
 45 Leu Leu Ala Ala Ser Pro Gly Gly Ala Leu Ala Arg Cys Pro Gly Cys
 46 20 25 30
 48 ggg caa ggg gtg cag gcg ggt tgt cca ggg ggc tgc gtg gag gag gag 143
 49 Gly Gln Gly Val Gln Ala Gly Cys Pro Gly Gly Cys Val Glu Glu Glu
 50 35 40 45
 52 gat ggg ggg tcg cca gcc gag ggc tgc gcg gaa gct gag ggc tgt ctc 191
 53 Asp Gly Gly Ser Pro Ala Glu Gly Cys Ala Glu Ala Glu Gly Cys Leu
 54 50 55 60
 56 agg agg gag ggg cag gag tgc ggg gtc tac acc cct aac tgc gcc cca 239
 57 Arg Arg Glu Gly Gln Glu Cys Gly Val Tyr Thr Pro Asn Cys Ala Pro
 58 65 70 75
 60 gga ctg cag tgc cat ccg ccc aag gac gac gag gcg cct ttg cgg gcg 287
 61 Gly Leu Gln Cys His Pro Pro Lys Asp Asp Glu Ala Pro Leu Arg Ala
 62 80 85 90 95

RAW SEQUENCE LISTING

DATE: 03/05/2004

PATENT APPLICATION: US/09/786,377A

TIME: 12:05:41

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03052004\I786377A.raw

```

64 ctg ctg ctc ggc cga ggc cgc tgc ctt ccg gcc cgc gcg cct gct gtt      335
65 Leu Leu Leu Gly Arg Gly Arg Cys Leu Pro Ala Arg Ala Pro Ala Val
66      100      105      110
68 gca gag gag aat cct aag gag agt aaa ccc caa gca ggc act gcc cgc      383
69 Ala Glu Glu Asn Pro Lys Glu Ser Lys Pro Gln Ala Gly Thr Ala Arg
70      115      120      125
72 cca cag gat gtg aac cgc aga gac caa cag agg aat cca ggc acc tct      431
73 Pro Gln Asp Val Asn Arg Arg Asp Gln Gln Arg Asn Pro Gly Thr Ser
74      130      135      140
76 acc acg ccc tcc cag ccc aat tct gcg ggt gtc caa gac act gag atg      479
77 Thr Thr Pro Ser Gln Pro Asn Ser Ala Gly Val Gln Asp Thr Glu Met
78      145      150      155
80 ggc cca tgc cgt aga cat ctg gac tca gtg ctg cag caa ctc cag act      527
81 Gly Pro Cys Arg Arg His Leu Asp Ser Val Leu Gln Gln Leu Gln Thr
82 160      165      170      175
84 gag gtc tac cga ggg gct caa aca ctc tac gtg ccc aat tgt gac cat      575
85 Glu Val Tyr Arg Gly Ala Gln Thr Leu Tyr Val Pro Asn Cys Asp His
86      180      185      190
88 cga ggc ttc tac cgg aag cgg cag tgc cgc tcc tcc cag ggg cag cgc      623
89 Arg Gly Phe Tyr Arg Lys Arg Gln Cys Arg Ser Ser Gln Gly Gln Arg
90      195      200      205
92 cga ggt ccc tgc tgg tgt gtg gat cgg atg ggc aag tcc ctg cca ggg      671
93 Arg Gly Pro Cys Trp Cys Val Asp Arg Met Gly Lys Ser Leu Pro Gly
94      210      215      220
96 tct cca gat ggc aat gga agc tcc tcc tgc ccc act ggg agt agc ggc      719
97 Ser Pro Asp Gly Asn Gly Ser Ser Ser Cys Pro Thr Gly Ser Ser Gly
98      225      230      235
100 taaagctggg ggatagaggg gctgcagggc cactggaagg aacatggagc tgtcatcact      779
102 caacaaaaaa ccgagccct caatccacct tcaggccccc ccccatgggc ccctcaccgc      839
104 tggttgga aaa gagtggtgt gttggctggg gtgtcaataa agctgtgctt ggggtcgctg      899
106 aaaaaaaaaa aaaaaaaaaa      918
109 <210> SEQ ID NO: 2
110 <211> LENGTH: 239
111 <212> TYPE: PRT
112 <213> ORGANISM: Homo sapiens
114 <400> SEQUENCE: 2
116 Cys Pro His Arg Leu Leu Pro Pro Leu Leu Leu Leu Leu Ala Leu Leu
117 1      5      10      15
120 Leu Ala Ala Ser Pro Gly Gly Ala Leu Ala Arg Cys Pro Gly Cys Gly
121      20      25      30
124 Gln Gly Val Gln Ala Gly Cys Pro Gly Gly Cys Val Glu Glu Glu Asp
125      35      40      45
128 Gly Gly Ser Pro Ala Glu Gly Cys Ala Glu Ala Glu Gly Cys Leu Arg
129      50      55      60
132 Arg Glu Gly Gln Glu Cys Gly Val Tyr Thr Pro Asn Cys Ala Pro Gly
133 65      70      75      80
136 Leu Gln Cys His Pro Pro Lys Asp Asp Glu Ala Pro Leu Arg Ala Leu
137      85      90      95
140 Leu Leu Gly Arg Gly Arg Cys Leu Pro Ala Arg Ala Pro Ala Val Ala

```

RAW SEQUENCE LISTING

DATE: 03/05/2004

PATENT APPLICATION: US/09/786,377A

TIME: 12:05:41

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03052004\I786377A.raw

```

141          100          105          110
144 Glu Glu Asn Pro Lys Glu Ser Lys Pro Gln Ala Gly Thr Ala Arg Pro
145          115          120          125
148 Gln Asp Val Asn Arg Arg Asp Gln Gln Arg Asn Pro Gly Thr Ser Thr
149          130          135          140
152 Thr Pro Ser Gln Pro Asn Ser Ala Gly Val Gln Asp Thr Glu Met Gly
153 145          150          155          160
156 Pro Cys Arg Arg His Leu Asp Ser Val Leu Gln Gln Leu Gln Thr Glu
157          165          170          175
160 Val Tyr Arg Gly Ala Gln Thr Leu Tyr Val Pro Asn Cys Asp His Arg
161          180          185          190
164 Gly Phe Tyr Arg Lys Arg Gln Cys Arg Ser Ser Gln Gly Gln Arg Arg
165          195          200          205
168 Gly Pro Cys Trp Cys Val Asp Arg Met Gly Lys Ser Leu Pro Gly Ser
169          210          215          220
172 Pro Asp Gly Asn Gly Ser Ser Ser Cys Pro Thr Gly Ser Ser Gly
173 225          230          235
176 <210> SEQ ID NO: 3
177 <211> LENGTH: 1196
178 <212> TYPE: DNA
179 <213> ORGANISM: Renilla reniformis
181 <220> FEATURE:
182 <221> NAME/KEY: CDS
183 <222> LOCATION: (10)..(945)
184 <223> OTHER INFORMATION:

```

W--> 187 <400> 3

```

188 agcttaaag atg act tcg aaa gtt tat gat cca gaa caa agg aaa cgg atg      51
189          Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met
190          1          5          10
192 ata act ggt ccg cag tgg tgg gcc aga tgt aaa caa atg aat gtt ctt      99
193 Ile Thr Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu
194 15          20          25          30
196 gat tca ttt att aat tat tat gat tca gaa aaa cat gca gaa aat gct      147
197 Asp Ser Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala
198          35          40          45
200 gtt att ttt tta cat ggt aac gcg gcc tct tct tat tta tgg cga cat      195
201 Val Ile Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His
202          50          55          60
204 gtt gtg cca cat att gag cca gta gcg cgg tgt att ata cca gat ctt      243
205 Val Val Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu
206          65          70          75
208 att ggt atg ggc aaa tca ggc aaa tct ggt aat ggt tct tat agg tta      291
209 Ile Gly Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu
210          80          85          90
212 ctt gat cat tac aaa tat ctt act gca tgg ttt gaa ctt ctt aat tta      339
213 Leu Asp His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu
214 95          100          105          110
216 cca aag aag atc att ttt gtc ggc cat gat tgg ggt gct tgt ttg gca      387
217 Pro Lys Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala

```

RAW SEQUENCE LISTING

DATE: 03/05/2004

PATENT APPLICATION: US/09/786,377A

TIME: 12:05:41

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03052004\I786377A.raw

```

218          115          120          125
220 ttt cat tat agc tat gag cat caa gat aag atc aaa gca ata gtt cac      435
221 Phe His Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val His
222          130          135          140
224 gct gaa agt gta gta gat gtg att gaa tca tgg gat gaa tgg cct gat      483
225 Ala Glu Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp
226          145          150          155
228 att gaa gaa gat att gcg ttg atc aaa tct gaa gaa gga gaa aaa atg      531
229 Ile Glu Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met
230          160          165          170
232 gtt ttg gag aat aac ttc ttc gtg gaa acc atg ttg cca tca aaa atc      579
233 Val Leu Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile
234 175          180          185          190
236 atg aga aag tta gaa cca gaa gaa ttt gca gca tat ctt gaa cca ttc      627
237 Met Arg Lys Leu Glu Pro Glu Glu Phe Ala Ala Tyr Leu Glu Pro Phe
238          195          200          205
240 aaa gag aaa ggt gaa gtt cgt cgt cca aca tta tca tgg cct cgt gaa      675
241 Lys Glu Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu
242          210          215          220
244 atc ccg tta gta aaa ggt ggt aaa cct gac gtt gta caa att gtt agg      723
245 Ile Pro Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg
246          225          230          235
248 aat tat aat gct tat cta cgt gca agt gat gat tta cca aaa atg ttt      771
249 Asn Tyr Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe
250          240          245          250
252 att gaa tcg gat cca gga ttc ttt tcc aat gct att gtt gaa ggc gcc      819
253 Ile Glu Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala
254 255          260          265          270
256 aag aag ttt cct aat act gaa ttt gtc aaa gta aaa ggt ctt cat ttt      867
257 Lys Lys Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe
258          275          280          285
260 tcg caa gaa gat gca cct gat gaa atg gga aaa tat atc aaa tcg ttc      915
261 Ser Gln Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe
262          290          295          300
264 gtt gag cga gtt ctc aaa aat gaa caa taa ttacttttggt tttttattta      965
265 Val Glu Arg Val Leu Lys Asn Glu Gln
266          305          310
268 catttttccc gggtttaata atataaatgt cattttcaac aattttatatt taactgaata 1025
270 tttcacaggg aacattcata tatgttgatt aatttagctc gaactttact ctgtcatatc 1085
272 attttggaat attacctctt tcaatgaaac tttataaaca gtggttcaat taattaatat 1145
274 atattataat tacatttggt atgtaataaa ctcggtttta ttataaaaaa a      1196
277 <210> SEQ ID NO: 4
278 <211> LENGTH: 311
279 <212> TYPE: PRT
280 <213> ORGANISM: Renilla reniformis
282 <400> SEQUENCE: 4
284 Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met Ile Thr
285 1          5          10          15
288 Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu Asp Ser

```

RAW SEQUENCE LISTING

DATE: 03/05/2004

PATENT APPLICATION: US/09/786,377A

TIME: 12:05:41

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03052004\I786377A.raw

```

289          20          25          30
292 Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala Val Ile
293          35          40          45
296 Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His Val Val
297          50          55          60
300 Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu Ile Gly
301 65          70          75          80
304 Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu Leu Asp
305          85          90          95
308 His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu Pro Lys
309          100          105          110
312 Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala Phe His
313          115          120          125
316 Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val His Ala Glu
317          130          135          140
320 Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp Ile Glu
321 145          150          155          160
324 Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met Val Leu
325          165          170          175
328 Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile Met Arg
329          180          185          190
332 Lys Leu Glu Pro Glu Glu Phe Ala Ala Tyr Leu Glu Pro Phe Lys Glu
333          195          200          205
336 Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu Ile Pro
337          210          215          220
340 Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg Asn Tyr
341 225          230          235          240
344 Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe Ile Glu
345          245          250          255
348 Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala Lys Lys
349          260          265          270
352 Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe Ser Gln
353          275          280          285
356 Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe Val Glu
357          290          295          300
360 Arg Val Leu Lys Asn Glu Gln
361 305          310

```

364 <210> SEQ ID NO: 5

365 <211> LENGTH: 543

366 <212> TYPE: DNA

367 <213> ORGANISM: Homo sapiens

369 <220> FEATURE:

370 <221> NAME/KEY: CDS

371 <222> LOCATION: (1)..(543)

372 <223> OTHER INFORMATION:

W--> 375 <400> 5

```

376 atg gga atc cca atg ggg aag tcg atg ctg gtg ctt ctc acc ttc ttg
377 Met Gly Ile Pro Met Gly Lys Ser Met Leu Val Leu Leu Thr Phe Leu
378 1          5          10          15

```

48

VERIFICATION SUMMARY

DATE: 03/05/2004

PATENT APPLICATION: US/09/786,377A

TIME: 12:05:42

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03052004\I786377A.raw

L:39 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:36
L:187 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:184
L:375 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:372
L:637 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:9,Line#:634